

Date Mailed: June 25, 2001

P.O. Box 5
Raleigh, NC 27602
(919) 854-1844

TRIGGERED PACKET DATA RATE CHANGE IN A COMMUNICATION SYSTEM

BACKGROUND OF THE INVENTION

The present invention relates to wireless communication systems in general, and particularly relates to such systems that employ variable rate or variable bandwidth radio channels.

As wireless communication systems evolve, the range of services offered by such systems expands. The increasing diversity, sophistication, and quality of available services have attracted a rapidly growing number of people who use wireless communications in their everyday lives. For example, today's mobile terminal user may subscribe to a wide range of voice and data services, including Internet access, email, and other data services in addition to more conventional voice services. It is expected that the demand for wireless communication services will continue to grow as new services are added and existing services are improved.

While the demand for data services continues to grow, the resources available to provide such services are finite. These resources include both the call processing and management resources, generally referred to as network resources, as well as the radio resources necessary to communicate with individual mobile terminals.

In next generation communication systems, the number of channels or bandwidth allocated to a user varies, depending on the services being provided to that user. For example, in cdma2000-based (IS-2000) networks, a user might be allocated a fundamental channel to support low data rate services, such as voice. For higher data rate services, the user might be allocated a fundamental channel and one or more supplemental channels to support the higher data rates. Thus, in contrast to older communication systems, these newer communication networks may allocate substantial resources to a single user, potentially reducing the total number of users that may be simultaneously supported.

SUMMARY OF THE INVENTION

The present invention comprises a method and apparatus for managing radio channel allocation and configuration in a radio access network. In one aspect, the present invention provides rate management functions that dynamically adjust the data rate capacity of an allocated radio channel upwards or downwards to its most appropriate rate setting. The appropriateness of the current data rate setting may be evaluated in terms of channel utilization, observed error, or any combination thereof. By selecting the best matching of available data rate settings for allocated radio channels, the radio access network's resources are more efficiently utilized. So, rather than maintaining a high data rate capacity on a channel that is being underutilized or suffers from excessive error, that channel's data rate may be reduced and the freed resources put to better use elsewhere.

Radio channel management may utilize a number of variables intended to balance the desire to maintain the minimal resource allocation for a given channel against the need to limit additional processing and signaling overhead associated with reducing or increasing radio channel capacity. These variables may include timing parameters that permit tracking the time since a last channel adjustment was made upwards or downwards, and may be used to prevent the rate management function from implementing rate control changes too frequently. Additional variables used in the dynamic rate management and radio channel allocation techniques of the present invention may further include subscriber-specific values, such as might be taken from home location register information. Thus, if a given subscriber has paid for premium access, higher access rates may be assigned or maintained for that subscriber.

Radio channel management in accordance with the present invention may be practiced in any number of radio network types, with cdma2000 and Wideband CDMA

(W-CDMA) being primary examples. In these types of radio networks, individual subscribers may be allocated one or more radio channels or one or more channel types or data rate capacities. By incorporating the rate management techniques of the present invention, the radio channel resources allocated to individual subscribers may be maintained at or near the actual capacity needed by each subscriber.

The present invention may also incorporate channel allocation techniques that determine the type or number of radio channels initially allocated to a given subscriber by the radio access network. These techniques may find particular use in radio access networks where subscribers may be initially allocated a basic rate channel, such as a fundamental channel in a cdma2000 network, which is then altered or supplemented with additional radio channel capacity if needed. Thus, given subscribers may be allocated additional radio channel data rate capacity based on actual need. Then, these higher data rate channel resources may be dynamically rate managed as above.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a diagram of an exemplary wireless communication network.

Fig. 2 is a diagram of typical RF channels allocated between a mobile terminal and the radio access network of Fig. 1.

Fig. 3 is a diagram of exemplary logic for high data rate channel activation and rate management activation.

Fig. 4 is a more detailed diagram of exemplary logic for high data rate channel activation.

Fig. 5 is a more detailed diagram of exemplary logic for radio channel data rate management.

Fig. 6 is an exemplary diagram for the up-rate management operations shown in Fig. 5.

Fig. 7 is an exemplary diagram for the down-rate management operations shown in Fig. 5.

DETAILED DESCRIPTION OF THE INVENTION

The present invention may be applied to a variety of radio access network types.

5 These network types include but are not limited to so-called third generation (3G) wireless cellular communication networks, such as those radio access networks based on the cdma2000 or W-CDMA standards. Thus, it should be understood that the following discussion of particular system types and implementations is only for illustrative purposes. It should also be understood that where illustrative variables are involved,
10 such as various timers or counters, and minimum and maximum values, these value are configurable as specific implementations might require.

Turning now to the drawings, Fig. 1 illustrates a wireless communication network generally referred to the by numeral 10. The network 10 comprises a radio access network (RAN) 12, and a core network comprising a mobile switching center (MSC) 14
15 and associated home location register (HLR) 16, and a packet data serving node (PDSN) 18. The network 10 provides communication between one or more mobile terminals 20 and one or more external communication networks, such as the Public Switched Telephone Network (PSTN) 22 and one or more packet data networks (PDNs) 24, such as the Internet.

20 The RAN 12 supports communication with the mobile terminals 20 via RF signaling. A typical RAN 12 comprises one or more base station controllers (BSCs) 30 and a number of associated radio base stations (RBSs) 32. The BSC 30 provides connection management services, communication signal processing, and radio link management functions. Operating under control of the BSC 30, the RBS 32 provides
25 the radio resources used to establish and maintain communication between the mobile

terminals 20 and the RAN 12. Typically, radio resources are allocated to a given mobile terminal 20 in the form of one or more radio channels.

Within this discussion, the term radio channel generally denotes the physical signaling resources allocated to establish a dedicated data or control link to a given mobile terminal 20. For example, if the RAN 12 uses a code division multiple access (CDMA) air interface, a radio channel includes the RF transceiver resources necessary to generate and receive signals to and from the mobile terminal 20, along with the unique spreading codes necessary to operate within the multiple-access environment inherent in CDMA systems. Walsh codes, which are orthogonal code sequences, are commonly used as CDMA spreading codes, with one Walsh code used to uniquely encode each defined radio channel.

In addition to managing the radio resources, the BSC 30 also configures and allocates network resources in association with its call management functions. As such the BSC 30 cooperates with the MSC 14 in call setup and teardown operations, as well as cooperating with the PDSN 18 to establish needed network channels to support packet data communications with the PDN 24. Call processing and management functions provided by the BSC 30 further include configuring and maintaining backhaul connections between it and the RBSs 32.

In operation, then, the BSC 30 allocates logical and physical network and radio resources as needed to the various mobile terminals 20 in active communication with the RAN 12. By properly allocating these resources, the BSC 30 supports and manages communication traffic to and from the mobile terminals 20. This traffic may involve a user of one mobile terminal 20 communicating with another mobile terminal user, or may involve the mobile terminal user communicating with other users or systems associated with the PSTN 22 or PDN 24.

The particular radio and network communication channels that are configured for a given mobile terminal 20 depend on the service or services required by the user of that mobile terminal 20. Fig. 2 illustrates the dedicated physical radio channels used for communications between the RAN 12 and the mobile terminal 20 in a cdma2000-based network. These channels may include a fundamental channel (FCH), a supplemental channel (SCH), and a dedicated control channel (DCCH). Forward channels carry data from the RAN 12 to the mobile terminal 20, while reverse channels carry data from the mobile terminal 20 to the RAN 12.

Fundamental channels are typically used for low data rate (e.g., 9.6 kbps) services such as voice, while the supplemental channels are used to support higher data rate services. Typically, the maximum data rate of a supplemental channel is configurable between some minimum and maximum value. For example, a typical forward SCH (F-SCH) in a cdma2000 network may support data rates from 9.6 kilobits-per-second (kbps) to 2.3 Mbps (mega-bits-per-second). If the RAN 12 is based on W-CDMA the terminology changes, as that standard does not use fundamental and supplemental channels per se, but nonetheless W-CDMA employs configurable or variable data rate radio channels.

In a typical cdma2000 scenario, a user of the mobile terminal 20 is assigned a forward FCH (F-FCH) upon initially accessing the RAN 12. If the user requires a higher data rate service, the user's mobile terminal may be assigned a F-SCH to support the higher data rate needed. In packet data calls, the F-FCH is setup during call-setup time. With higher rates of incoming communication traffic (i.e., more than supportable with F-FCH capacity), the radio link protocol (RLP) data queue used by the RAN 12 to queue outgoing communication traffic to the mobile terminal 20 will grow in length. This leads to increasing queuing delay, which beyond a given delay threshold is undesirable.

To understand the relationship between queuing delay and queue length, first consider end-to-end delay, also referred to as round trip time (RTT) of the data connection involved. In this exemplary case, the RTT may be on the order of 500 ms to one second, which equals the typical packet data Transmission Control Protocol (TCP) time out. The wireless link portion of the RTT delay varies depending upon the air interface, but may typically range from 80 to 120 ms. Also, a nominal frame error rate (FER) of roughly two-percent may be assumed for the outgoing communication traffic. With these values, the effective maximum wireless link on the F-FCH is given approximately as,

Rate = (base FCH rate)(% of available bandwidth)(1 – FER),
which approximately equals $9.6 \text{ kbps} * 0.8 * 0.98 = 7.526 \text{ kbps}$.

Thus, the effective RLP buffer length (data queue) permissible to avoid delays in excess of a target value (e.g., 300 ms in this particular example) is given as,

$$\text{Maximum buffer length} = 7.526 \text{ kbps} / 8 \text{ bits/byte} * 0.3 \text{ seconds} = 282 \text{ bytes}.$$

Therefore, in this example, if the F-FCH RLP data queue for outgoing communication traffic to the mobile terminal exceeds 282 bytes excessive transmission delay may occur.

It follows from this data queue length analysis that excessive packet size can violate the queue length restraint. That is, the communication traffic may comprise packet data of varying packet sizes. The size of a given packet depends on a number of variables, including its type and source. In any case, in the above example, a large data packet would violate the maximum desirable queue length, and could result in undesired transmit delay from the perspective of the mobile terminal.

However, the packet size limit above which supplemental channel activation is triggered may or may not be equal to the queue length limit. For example, some types of communication traffic may comprise occasional large data packets with relatively long

intervening periods of no traffic. Wireless application protocol (WAP) traffic may exhibit this characteristic. As such, it would be undesirable to trigger F-SCH allocation just to service the occasional large, intermittent data packet. Thus, the trigger threshold should be configurable and may take on one of several different values that may be set based on traffic source, traffic type, or other operating condition.

One approach to determining whether to assign a F-SCH to a given mobile terminal 20 is outlined in the logic flow of Fig. 3. In Fig. 3, processing begins (block 200) assuming that a F-FCH is already assigned to mobile terminal 20 and that some form of data rate monitoring indicates the need for radio channel capacity beyond that provided by the F-FCH. With regard to the above discussion, this may be, for example, an indication that the RLP queue length for outgoing communication traffic is excessive, or that packet size is larger than expected. If a F-SCH is not already assigned to the mobile terminal 20 (block 202), a F-SCH activation timer is incremented (block 204). If the activation timer has reached a configurable, predetermined value (block 206), it is reset and procedures to selectively allocate a F-SCH to address the needed data rate increases are called (block 208). Exemplary details regarding F-SCH activation appear later in Fig. 4.

If the activation timer has not reached the predetermined value (block 206), processing ends (block 210). These processing steps may be called with a frequency dependent on the desired maximum transmit delay time. Thus, the logic of Fig. 3 may be executed every 60 ms for example. With that value, the activation timer would reach its threshold value after $n * 60$ ms, where "n" is the configurable threshold count value.

If a F-SCH is already active for the mobile terminal 20 (block 202), then rate management techniques may be used to maintain an optimal or best matching data rate capacity configuration for the assigned F-SCH resources. Such rate management functions should be called periodically during active service, so a rate change timer is

incremented each time the logic of Fig. 3 is called with an active F-SCH (block 212).

Again, an exemplary calling interval might be 60 ms, or other appropriately selected period. If the change timer reaches a predetermined count value "m" (block 214), a rate management routine is called and the rate change timer is reset (block 216). Rate management features are detailed later herein. Upon completion of rate management operations or if the rate change timer has not reached its threshold count, processing ends (block 210).

Fig. 4 provides exemplary flow logic for the F-SCH activation operations described above and shown in Fig. 3 (block 208). Processing begins (block 220) with a determination of whether F-SCH activation has been previously requested in association with the current mobile terminal communications session (e.g., the same call) (block 222). A request separation timer is used to prevent activation operations from requesting additional radio channel allocations (i.e., a new F-SCH assignment) too frequently. Without some mechanism to enforce a delay between requests, the channel activation process might repeatedly issue allocation requests, which must be processed at higher network layers, such as within the signaling layer, and could place significant, unproductive processing burden on the BSC 30.

If the request separation timer has a non-zero value, meaning that a prior channel activation request was issued that resulted in no F-SCH channel allocation, presumably due to a lack of available resources, the request separation timer is decremented (block 224) and processing ends (block 226). The activation routine will follow this logical flow through subsequent calls to it until the request separation timer has been decremented to zero, thus enforcing the desired delay between channel activation requests sent to the signaling layer. One second represents an exemplary value for the request separation timer, and with this value repeated requests for SCH

activation would be at least one second apart, thereby minimizing processing overhead in the RAN 12.

If the request separation timer is already zero, then one or more F-SCH activation "triggers" are evaluated. In this exemplary logic flow, the first trigger condition evaluated is the length of the RLP data queue for the outgoing communication traffic to the mobile terminal 20 (block 228). If the outgoing packet data has an excessive queue length, this is an indication that the F-FCH is inadequate to serve the mobile terminal's service needs. Simply put, the queue length of outgoing data is an effective indicator as to whether the data rate of the F-FCH is sufficient to keep up with the incoming data rate of traffic for the mobile terminal 20. Here, "incoming" refers to traffic coming into the RAN 12 for transmission to the mobile terminal 20.

If the RLP data queue length does not exceed a defined threshold (e.g., a maximum permissible queue length), an excess queue length timer is reset (block 230), and the next trigger condition is evaluated. The next trigger is based on the size of one or more packets in the outgoing communication traffic. For example, if one or more packets have a packet size that exceeds the maximum permissible packet size, then activation of the F-SCH may be warranted to avoid excessive transmission delay to the mobile terminal 20 (block 232).

If the current RLP data queue length exceeds maximum permissible queue length value (block 228), the excess queue length timer is incremented (block 234). An excess queue length timer at or above a defined threshold count indicates that the RLP data queue has exceeded its maximum permissible length for a defined period of time and, therefore, activation of the F-SCH is warranted (block 232). Note that the interval between calling the processing logic of Fig. 4 will determine the amount of time represented by each increment of the excess queue length timer, with a value of 60 ms representing a nominal time period. Using the timer insures that the excess queue

length trigger is based on a continuing condition rather than on a temporary increase in RLP data queue length. If either the excess RLP data queue length or excess packet size trigger conditions are satisfied (block 232), processing continues with F-SCH activation (block 236).

5 A number of steps occur with regard to the F-SCH activation request. First, an activation request rate value (Req_Rate) is set to the desired data rate capacity indicator. The requested rate may default to, for example, the maximum defined data rate, or may be based on some intermediate value that reflects a rate that is satisfactory for most general high data rate traffic. Alternatively, the requested rate may be based on
10 a custom value that is linked to subscriber-specific information. For example, the user of the mobile terminal 20 may have relevant subscriber profile information in the HLR 16 (or elsewhere) that determines the default data rate. This subscriber information may indicate, for example, whether the subscriber is eligible for premium data rate services, or may otherwise have a default data rate value that may be used in requesting the F-
15 SCH.

 An activation procedure is called using the Req_Rate value that involves communicating with other processes in the BSC 30 where the resource assignments are made, such as communicating with the signaling layer. If the requested assignment is granted, the assigning process communicates back an Assigned_Rate value that may or
20 may not equal the Req_Rate value. An Initial_Rate value is set equal to the Assigned_Rate so that in indicator of the beginning data rate for the assigned F-SCH may be preserved. An FSCH_Active flag may then be set to indicate that F-SCH resources are active for the mobile 20, and an adjustment guard timer (Guard_Timer) may be initialized with an appropriate count value. The Guard_Timer is described later
25 in the discussion of rate management operations. After completion of these operations, processing ends (block 226).

Note that if the BSC 30 indicates that no additional resources are available for completing the requested radio channel allocation, the request separation timer may be reset to an appropriate count value (block 236). As explained, subsequent calls to the logic flow of Fig. 4 will then decrement that timer, thus separating repeated resource requests by an appropriate delay to limit system processing overhead.

While the above discussion of F-SCH activation triggering related primarily to the cdma2000 scenario where an initial F-FCH is supplemented with a F-SCH if needed, it alluded to rate management processing once the F-SCH channel was allocated. One advantage of rate management is that it allows the RAN 12 to avoid under- or over-allocating radio resources. That is, the data rate capacity allocated to individual mobile terminals 20 should ideally match the actual data rate requirements of those mobile terminals 20.

While exact data rate matching may not be practicable since rates may be selectable only in discrete steps and user traffic is very bursty, it is still preferable to assign the most appropriate data rate capacity to each user. It should be noted that data rate management as described herein is not limited to cdma2000, or to W-CDMA. Data rate management as described herein is appropriate for wireless communication systems where multiple data rates are defined and radio resource allocation may be optimized by matching allocated radio channel rate capacity based on actual channel usage.

In general, an allocated radio channel is rate managed based on monitored channel usage or monitored error. This applies whether the channel in question is a F-SCH in a cdma2000 system, a configurable-rate radio channel, such as a digital traffic channel (DTCH), in a W-CDMA system, or other configurable rate channel in another type of communication system.

Channel usage and channel errors are generally good bases for rate management functions. In a cdma2000 implementation, the RAN 12 (e.g., BSC 30) may monitor RLP usage and RLP FER. Other network types may use different terminology for like parameters. Thus, if RLP usage is above a configurable upper usage threshold, it may be appropriate to re-configure the allocated radio channel to provide a higher data rate capacity to the mobile terminal 20. Conversely, if RLP usage is below a configurable lower threshold, it may be appropriate to re-configure the allocated radio channel to provide a lower data rate capacity, or perhaps even de-allocate the radio channel.

RLP usage (or its comparable metric) may be measured in exemplary fashion as follows,

$$\text{RLP_Usage} = (X * 100) / (D * \text{RATE}),$$

where X = total number of octets sent as new data to the mobile terminal 20, D = number of frame durations since last usage measurement, and RATE = a value based on the current channel data rate. Essentially, RLP_Usage serves as an indicator of the percentage of RLP frames used to send new data to the mobile terminal 20.

As with usage, the capacity (i.e., data rate) of the allocated radio channel may be adjusted upwards or downwards based on observed communication error, which may be indicated by RLP FER for example. A high FER indicates a potentially bad radio link with the mobile terminal 20 and suggests that forward link power levels are not sufficient to sustain communications at the current data rate.

RLP error (or its comparable metric) may be measured in exemplary fashion as follows,

$$\text{RLP_Error} = (E * 100) / (D * \text{RATE}),$$

where E = total number of retransmitted octets. Essentially, RLP_Error serves as an indicator of the percentage of RLP frames being used to retransmit data to the mobile terminal 20.

Using these usage and error parameters, rate management processing may
5 adjust the capacity of allocated radio channels as needed. The frequency at which rate changes are made may be adjusted in accordance with particular system needs. For example, in a system with a relatively slow forward-link power control loop, it may be desirable to slow down rate adjustments such that rate change adjustments do not outpace the frequency with which forward link power adjustments are made. Such a
10 scenario might be found in, for example, a wireless local loop (WLL) implementation of the RAN 12, where mobility is typically not an issue and power control is typically implemented with a much slower control bandwidth as compared to a conventional high-mobility cellular communications implementation of the RAN 12.

Fig. 5 illustrates exemplary logic for a rate management function. The rate
15 management function determines whether to increase, decrease, or maintain a current data rate for a radio channel allocated to the mobile terminal 20. Typically, a higher-level program or other controlling function calls the rate management function on a periodic basis. In an exemplary implementation with a data frame rate of 20 ms, the rate management function may be called every six frames or every 120 ms.

20 Upon being called, rate management processing begins (block 250) with a first or one-time initialization of various timers and constants as needed. Here, High and Low Duration Timers are initialized to zero. As the rate management function is called on a periodic basis, the duration of the calling interval determines High and Low Duration Timer values. That is, if the rate management function is called once per N frame cycles
25 (i.e., once per N data frames), each count value of the High and Low Timers represents

N frame times. With the exemplary value of six (6) for N, each count value of the Low and High Duration Timers equates to 120 ms for 20 ms frame times.

Processing continues with evaluation of the Guard Timer (block 252). The Guard Timer prevents the rate management function from making rate adjustments too frequently. That is, after each rate adjustment, it is reset to a default non-zero count and must decrement down to zero again before a subsequent rate change is permitted. If the Guard Timer is not clear (i.e., non-zero), it is decremented (block 254) and processing ends (block 256). Note that here, as in the other logic flow illustrations, "RETURN" indicates a return to other processing, which most likely entails a return to the calling program for execution of other tasks.

If the Guard Timer is clear (i.e., zero). Processing continues with evaluation of the Up Delay Timer (block 258), which operates on a similar basis, but may have a different non-zero default or reset count value than the Guard Timer. The Up Delay Timer prevents a too rapid rate increase following a rate decrease operation and will be explained in more detail later.

If the Up Delay Timer is not clear (i.e., non-zero), it is decremented (block 260) and rate management processing ends (block 256) until the next time it is called. If the Up Delay Timer is clear, processing continues with an evaluation of the channel usage (block 262). An exemplary channel usage evaluation compares the RLP_Usage value with a MAX_USAGE constant, which may be a configurable percentage utilization threshold. The calculation of RLP_Usage was described earlier, and it was explained that it indicated a percent utilization of the allocated radio channel as determined in the time between calls to the rate management function.

If channel utilization is greater than MAX_USAGE, the High Duration Timer is incremented and the Low Duration Timer is set to zero (block 264). If the High Duration Timer is satisfied, meaning it has reached a predetermined count value (block 266), up-

rate management operations are performed at "A," which are detailed in Fig. 6, and processing ends (block 256). Use of the High Duration Timer, which is incremented each time the rate management function is called if channel usage (RLP_Usage) is above the MAX_USAGE threshold, serves to qualify the high usage condition. That is, if the high usage condition has persisted over a sufficient number of calling intervals to the rate management function, increasing the data rate capacity to the mobile terminal 20 is warranted. Up-rate management operations are identified as "A" in Fig. 5, and detailed in Fig. 6. An interval count of three is an exemplary value for the High Duration Timer count threshold, based on a maximum 60 ms frame timing.

If channel utilization is not greater than MAX_USAGE (block 262), the High Duration Timer is set to zero (block 268). Processing continues with an evaluation of whether the current usage (RLP_Usage) is less than a usage limit for halving, or whether the channel errors (RLP_Error) exceed a defined error threshold denoted as MAX_ERROR (block 270). The usage limit for halving may be set depending on the type of RAN 12, but is generally defined as a minimum percentage of channel utilization below which a current data rate capacity should not be maintained.

Similarly, MAX_ERROR may be an FER threshold above which it makes sense to reduce the data rate to the mobile terminal 20. In other words, if the radio link conditions to the mobile terminal 20 are so bad that substantial errors occur at the current data rate, the allocated channel resources are not being efficiently used. In this sense, evaluation of channel utilization entails both the calculated percentage utilization, as well as the effectiveness of that utilization based on observed or measured error.

If neither the usage limit for halving nor the MAX_ERROR thresholds are violated, the Low Duration Timer is cleared to zero (block 272) and processing ends (block 256). However, if one or both thresholds are violated, the Low Duration Timer is incremented (block 274) and then evaluated to determine if the under-utilization

condition has persisted long enough to warrant a down-rate adjustment (block 276).

This condition is evaluated by determining whether the Low Duration Timer has reached a predetermined count threshold. If it has not, processing ends (block 256). If the Lower Duration Timer has reached the count threshold (e.g., a count of three), processing
5 continues with down-rate management operations at "B," which are detailed in Fig. 7.

In Fig. 6, the up-rate operations identified as "A" in Fig. 5 begin with an evaluation of whether the current data rate is less than the data rate of the radio channel in question when it was initially assigned to the mobile terminal 20 (block 280). If so, a requested rate parameter denoted as REQ_RATE is set to reflect a desired increase
10 from the current data rate. In this example, REQ_RATE is set to twice the current rate (block 282).

A request for a rate increase is then sent to the appropriate handler in the RAN
12 (e.g., the signaling layer functions in the BSC 30) using the REQ_RATE value to indicate the new rate desired (block 284). The signaling layer will allocate resources as
15 available and return a value indicating the assigned rate, which may or may not match the requested rate.

To preserve the rate actually assigned, this value is saved as the new current rate, denoted by CURRENT_RATE (block 286). Note that this logic flow assumes that resources to support a higher rate are available. However, the BSC 30 may not be able
20 to assign any higher rate and it should be understood that the logic of Fig. 6 implicitly accounts for this underlying possibility. Also, note that in some communication network types, increasing the data rate capacity may comprise assigning a new radio channel in addition to the currently assigned radio channel. As an example, an IS-2000 network may add supplemental channels (SCHs) as needed to accommodate the required data
25 rate. In any case, processing continues with resetting the Guard Timer (block 256), to

enforce a delay before any subsequent rate changes are made by the rate management function.

If the current rate is not less than the initial rate (block 280), processing continues with an optional evaluation of whether it is permissible to increase the rate beyond that which was initially allocated (block 290). If such an increase is not permitted, processing ends (block 256). However, if the increase is permissible, the REQ_RATE parameter is set to the desired rate value (block 292). The desired rate may be a multiple of the current rate, may be a default maximum value, or may be a custom value linked to, for example, subscriber-specific information. In this latter case, subscriber profile information, such as might be stored in the HLR 16, may be used to determine the value for REQ_RATE. After setting the REQ_RATE value, processing continues as above with the rate management function sending the rate request to the signaling layer and responding the assigned rate (blocks 284-288).

Allowing the current rate to progress above the initially assigned rate may have differing implications depending upon the type of RAN 12 involved. For example, in a cdma2000 system, the Walsh code assignments and the particular radio resources assigned to a given radio channel are based on the channel's data rate. To move beyond this initial rate may require assigning a new radio channel, which requires tearing down the existing channel and setting up the new channel. Moreover, there may be a period of overlapping teardown and setup where resources from both the old channel and new channel are dedicated to the mobile terminal 20. Thus, some care should be exercised in these operations, given the call processing overhead and resource utilization considerations involved.

With these considerations, it may be desirable to use custom subscriber information to determine whether a higher-than-initial data rate increase is permitted. The subscriber may be charged a premium for such flexibility, or may forego additional

service charges at his or her option. Incorporating subscriber preferences into the rate management function allows service providers to provide pricing and service flexibility to subscribers.

The remaining operations of the exemplary rate management function address
5 down-rate management shown in Fig. 5 as "B" and are illustrated in Fig. 7. Processing
begins with an evaluation of whether the current rate (CURRENT_RATE) is less than or
equal to the min rate, which is denoted as MIN_RATE, and whether current usage is
less than or equal to the minimum usage threshold MIN_USAGE, described earlier
(block 300). Here, MIN_RATE may be a value dependent on the type of RAN 12
10 involved, such as the minimum defined F-SCH data rate in a cdma2000 system, or the
next-to-lowest data rate defined for W-CDMA for example.

If the current rate is not greater than the minimum rate (block 303), processing
continues with no adjustment (block 256). If at or below MIN_RATE and at or below
MIN_USAGE, the higher data rate channel is deactivated, or a lower data rate channel is
15 otherwise adopted (block 302). In cdma2000, this might entail de-allocating the F-SCH,
relying only on the F-FCH. In W-CDMA, this might entail dropping back to the lowest
defined data rate for the allocated channel. Once the channel is de-allocated or a
minimum rate channel is adopted, processing continues with the Low Duration Timer
being cleared to zero (block 304), after which processing ends (block 256).

20 If either of the above conditions is not met, and if the current rate is greater than
the minimum rate, processing continues with additional evaluation of the channel usage
(block 306). A defined table of usage values may be stored in a memory comprising a
portion of the BSC 30 or elsewhere, such that it is accessible to the rate management
function. This table, denoted as Usage_Limit[], comprises a set of thresholds
25 corresponding to convenient fractional channel utilizations. For example, Usage_Limit[]
may comprise the following values:

Table 1: Usage_Limit Data Structure	
Usage_Limit[1/8]	1/8 Utilization Threshold
Usage_Limit[1/4]	1/4 Utilization Threshold
Usage_Limit[1/2]	1/2 Utilization Threshold

Of course, the definitions and thresholds in the above table may be varied as needed.

The first evaluation determines whether the channel usage (e.g., RLP_Usage) is below the Usage_Limit[1/8] threshold. If so, the REQ_RATE value is set to 1/8 the CURRENT_RATE value (block 308). This REQ_RATE value is checked to insure that is not below the minimum defined rate (block 310). If it is, the REQ_RATE is set to the minimum defined rate MIN_RATE (block 312) and processing continues. The current error level (e.g., RLP_Error) is evaluated to determine whether it exceeds the MAX_ERROR threshold (block 313). If the error threshold is exceeded, the Up Delay Timer is then set to enforce a delay that prevents subsequent calls to the rate management function from making upward rate adjustments too soon (block 326).

From there, processing continues with a request to the signaling layer to adjust the current rate to the newly requested reduced rate (block 314) and the Guard Timer is reset to enforce a delay before subsequent adjustments are made (block 316). Because reducing the current data rate addresses the low duration condition that prompted the rate fallback, the Low Duration Timer is reset to zero (block 304).

If the initial channel usage evaluation above determines that RLP usage was above the 1/8 threshold (block 306), processing continues with an evaluation of whether usage is below the 1/4 threshold (block 318). If so, the rate to be requested is set to 1/4 the current rate (block 320) and processing continues as above (e.g., from block 310). If the usage equals or exceeds the 1/4 threshold, the requested rate is set to 1/2 the current rate (block 324), and processing continues as above (e.g., from block 310).

If the condition necessitating downward rate adjustment was excessive error rather than underutilization of the channel, there is substantial value in imposing a

controlled delay using the Up Delay Timer before any subsequent up-rate adjustments are allowed. That is, it is likely that if the radio environment surrounding the mobile terminal 20 was bad enough to result in the FER exceeding the MAX_ERROR threshold, there will likely be little gained by bumping the data rate back up soon after a downward
5 adjustment.

Whether a downward rate adjustment is due to error or underutilization, it may also be advantageous to set the Up Delay Timer in consideration of the new rate. For example, if the new rate is at or near the highest defined data rate, it may be sensible to enforce a relatively long delay before allowing subsequent up-rate adjustments.

10 Conversely, if the new rate is near the minimum data rate, it may be sensible to impose a less lengthy delay before allowing subsequent up-rate adjustments. This logic acknowledges the increasing likelihood of failure in attempts to increase the data rate of a connection with an already high data rate.

Exemplary values for the Up Delay Timer in a cdma2000 environment are given
15 in the following table:

Table 2: Exemplary Up Delay Timer Settings	
New Rate = 16x (max rate setting)	Timer value = $16 * 8$
New Rate = 8x	Timer value = $8 * 8$
New Rate = 4x	Timer value = $4 * 8$
"x" denotes base data rate	

Of course, the up-rate delays are preferably configurable to accommodate the varying needs of different implementations and RAN types.

It should be understood that the preceding logic flows illustrated in Figs. 4-7 are
20 exemplary. The present invention may, of course, be carried out in other specific ways than those herein set forth without departing from the spirit and essential characteristics of the invention. The present embodiments are, therefore, to be considered in all

respects as illustrative and not restrictive, and all changes coming within the meaning and equivalency range of the appended claims are intended to be embraced therein.